

Classification with Random Forest Based on Local Tangent Space Alignment and Neighborhood Preserving Embedding for MSER features: MSER_DFT_LTSA-NPE_RF

Sevcan Aytaç Korkmaz*, Furkan Esmeray**

*Fırat University / Electronic and Automation Department, Elazığ, Turkey

**Munzur University / Electric and Energy Department, Tunceli, Turkey

Abstract— In this article, 180 gastric images taken with Light Microscope help are used. Maximally Stable Extremal Regions (MSER) features of the images for classification has been calculated. These MSER features have been applied Discrete Fourier Transform (DFT) method. High-dimensional of these MSER-DFT feature vectors is reduced to lower-dimensional with Local Tangent Space Alignment (LTSA) and Neighborhood Preserving Embedding (NPE). When size reduction process was done, properties in 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 dimensions have been obtained. These low-dimensional data are classified by Random Forest (RF) classification. Thus, MSER_DFT_LTSA-NPE_RF method for gastric histopathological images have been developed. Classification results obtained with these methods have been compared. According to the other methods, classification results for gastric histopathological images have been found to be higher.

Keywords—Maximally Stable Extremal Regions; Neighborhood Preserving Embedding; Local Tangent Space Alignment; Random Forest

I. INTRODUCTION

Stomach cancer is a type of cancer that occurs in the Stomach cancer, Stomach tissue and Stomach wall. According to studies conducted by the Ministry of Health, Stomach cancer was identified as thesecond most common cancer type. Endoscopy is the most important factor in the early diagnosis of this disease. Endoscopic examination of the endometrium and biopsy specimens are taken and diagnosed as pathological examinations. It is seen that half of the people who have this diseasea relate in the diagnosis and doctors cannot apply any treatment [1,2]. The most common sites of this disease in the world are distant countries such as Japan and China. Japan, the number of people with Stomach cancer accounts for about 30% of other cancer diseases. In the Americas, the number of Stomach cancer people increases every year [1,3-4]. According to research conducted worldwide, 26% of malesand 11% of females have Stomach cancer. Stomach cancer is located in the third place after lung and breast cancer in women and second place after lung cancer in males. According to the number of new Stomach cancer is estimated to be around 30 thousand a year [1,5]. S. Yoshihiro and colleagues [6] studied Stomach cancer by developing computer-based systems that can predict the risk factor. In the system developed, endoscopy images were taken from patients carrying H. pylori bacteria. 15 parameters were used to classify the Stomach

mucosa with 3 parameters on the back panel. The classification data were processed by Bayesian theorem and outputs were obtained. This study is the source for the treatment of patients who are at risk of Stomach cancers or patients who have to undergo endoscopy [1]. D.Ahmadzadeh [7] developed a cancer diagnosis system by using KDM (decision support machine) and local pattern algorithm methods for Stomach cancer diagnosis. By using the feature identification, feature extraction and noise reduction steps in the system they have developed, the results of estimation of 91,8% of 55 randomly selected samples were found. The common feature of both systems used in the study is that it is a system that helps the doctor in time and material sense [1]. Akbari et al. [8] developed a Stomach cancer diagnosis system by using infrared ultra-spectral imaging method. This study was performed by selecting patients with Stomach cancer. Spectral features were extracted from cancerous and normal tissues and compared with this, KDM method was used to determine the detection of cancerous regions by spectral diagram It was performed. High performance was obtained from the system by using HFD and Log transformation in the direction of the obtained data. In the study performed, 25 patients with indole cancer were detected in 30 patients and 83,3119% of the obtained system was mathematically successful [1]. Stomach cancer is a type of cancer that, when diagnosed late, leads the patient to death [1, 9].Stomach cancer usually begins

with ulcer and gastritis complaints. Cancer can affect lymph nodes and other peripheral organs [1,10].

Methods of the study are given in Section 2. Experiment results are described in Section 3. Discussions are described in Section 4, and conclusions are explained in Section 5.

The purpose of this article is to add a different study to the literature to help early diagnosis of stomach cancer using images of histopathology. The novelty of this study, when studies in the literature are examined, it is seen that MSER, DFT, LTSA, NPE and RF methods are not used together to help diagnose early gastric cancer. We have put forward a more powerful computer-aided method of using these methods together in the same moment. Also, the classification performances have been compared according to the number of selected MSER features.

II. THEORY AND METHOD

A. Maximally Stable Extremal Regions (MSER)

The MSER algorithm is an algorithm used to find circles or ellipse-like shapes (blobs) in images. The algorithm selects key points taking these shapes into account and calculates their attributes on these key points [11]. An MSER region [12] consists of a set of interconnected points over a certain threshold value, whose persistence is dependent on the changing threshold value. In other words, the selected region is a local binary form that does not depend on a set of threshold values. According to the study given in Reference [13], finding MSER regions works in a similar way to the Water shedding method. The threshold value is changed to vary between [0-255], while regions connected to one another, which do not change or change very little in all scenes, are called as MSER regions.

In the implementation of the method; the image points are sorted by the brightness value. and the points are placed in the image in ascending or descending order. During this process, the component regions and fields is kept in a list by finded with the join detection algorithm. At each step of the thresholding process, the small region with the association of the two regions is included in the large region. and the smaller region is removed from the list. The threshold values at which the variation of the zone area is minimized by increasing / decreasing the threshold value are selected as the threshold ranges producing the maximum stable extremal zones. In other words, the boundaries of the end regions formed by the connected components at all threshold values are represented

as $Q_1, \dots, Q_{i-1}, Q_i, \dots$ a series of contiguous regions. This sequence provides the $Q_i \subset Q_{i+1}$ condition. In order to select the Q_i^* extremal region of the array as the maximum stable [13]; To be local minimum is required in i^* value of the $q(i) = |Q_i \pm \Delta|/|Q_i|$ expression. In this expression, $| \cdot |$ denotes the area of the region, and Δ value denotes a parameter of the method. The \pm sign also indicates that there are local minimum for both decreasing and increasing threshold values. This process is applied on all the region arrays in the set of extremal regions to obtain the maximum extremal stable regions [13].

In the MSER method, neighboring image points having a similar color are subjected to clustering based on the stacked clustering. For clustering process, the color distances of four or eight neighboring points to one another are kept at an associated list. At each step of the algorithm, $t \in [0 \dots T]$, the image points are labeled progressively. If the coordinate space of the image points is denoted as a set of labels $\Omega = [1 \dots L] \times [1 \dots M] \subset Z^2$, each step is expressed as $E_t: \Omega \rightarrow N$ mapping. As a result of the labeling, the connected points with the same indicate specify the image extremal regions. The distance between all neighboring points of the image extremal region must be lower than $d_{thr}(t)$, which is a threshold value calculated for the step in question. The distance of the image points in the color space is calculated using the chi-square distance. Initially, all values in the E_0 label image are labeled as 0. In E_t tag image, all neighbor points with a distance less than $d_{thr}(t)$ are labeled as new region and E_{t+1} tag image is obtained [13]. Due to the spatial relationship between the image points, the distances of all neighboring points do not show a uniform distribution. The vast majority of distances have small values, and large distance values in very few number exist. Therefore, at each step, the threshold value is increased linearly, resulting in a very rapid number of tag changes at the beginning. Thus, to the end of the steps, the labels of many points change [13]. In order to change the label of an equal number of image points in each step, the distance between all neighboring points in the image is taken as a random variable and the threshold values are modified according to the ordered reversal of the cumulative distribution function (CDF) of this random variable [13]. The chi-square CDF for colored images is calculated as in Equation (1)

$$C_3(x) = \sqrt{\frac{4x}{\lambda\pi}} e^{-3x/2\mu} + \text{erf}(\sqrt{3x/2\mu}) \quad (1)$$

where μ is the mean of the sample set. As a result, Equation (2) is used to find the threshold values after the average estimate.

$$d_{thr}(t) = c^{-1} \left(\frac{t}{T} \right) \quad t \in [0 \dots T] \quad (2)$$

Next, field changes of the extremal regions detected by increasing the value of $d_{thr}(t)$ at each step are checked. and maximum stationary ones are detected. In addition, those smaller than a given value are eliminated from the maximum stable extremal regions [13].

B. Discrete Fourier Transform (DFT)

Fourier transformation is a mathematical method that allows a periodic signal to be expressed by sinus and cosine components at different frequencies [14]. The Fourier transformation is defined for an array of infinite lengths, and more importantly, a function of the w angular frequency, which is a continuous variable. When using MATLAB, we do not limit arrays and we need to evaluate for a limited number of points. Discrete Fourier Transform (DFT) eliminates these problems. In this article, a discrete Fourier transform (DFT) of multidimensional MSER property values is calculated using a Fast Fourier Transform (FFT) algorithm. The discrete Fourier transform is expressed as follows

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk} \quad 0 \leq k \leq N-1 \quad (3)$$

where is $W_N = e^{-j(\frac{2\pi}{N})}$. if the N sequence length is large, the direct presence of the DFT requires a large amount of processing. That is, as the N number increases, the number of transactions made increases rapidly and the number of transactions goes to an unacceptable level. In 1965, Cooley and Tukey developed a procedure to reduce the amount of processing required for Discrete Fourier Transform [14]. This procedure caused a sudden increase in DFT applications in digital signal processing and other fields. It has also been a pivotal step in the development of other algorithms. All these algorithms are known as Fast Fourier Transform (FFT) algorithms. These algorithms have greatly reduced the number of operations required for the DFT account, thereby ensuring ease of operation. FFT is an efficient and economical algorithm for DFT computation [15].

C. Local Tangent Space Alignment (LTSA)

Given A m -dimensional points sampled possibly with noise from an underlying d -dimensional manifold, this algorithm produces A dimensional coordinates $T \in R^{dxn}$ for the manifold constructed from f local nearest neighbors [16].

Step 1. Extracting local information:

For each $a = 1, \dots, A$, Determine f nearest neighbors x_{ab} of x_a , $b = 1, \dots, f$. Compute the d largest eigenvectors g_1, \dots, g_d of the correlation matrix $(X_a - \bar{x}_a e^T)^T (X_a - \bar{x}_a e^T)$, and set

$$G_a = [e/\sqrt{f}, g_1, \dots, g_d] \quad (4)$$

Step 2. Creation of the alignment matrix: If a direct eigenvalue decomposition is used, create the alignment matrix Φ according to the local collection. Otherwise, apply a routine that computes the matrix-vector multiplication for an arbitrary vector u [16].

Step 3. Calculate the spherical coordinates: Calculate the smallest eigenvector of $d+1$ and select the eigenvector matrix $[u_2, \dots, u_{d+1}]$ corresponding to the smallest eigenvalues, and $T = [u_2, \dots, u_{d+1}]^T$ [16].

D. Neighborhood Preserving Embedding (NPE)

Neighborhood Preserving Embedding (NPE) is a linear approach to the LLE algorithm [17]. The algorithmic procedure is formally described below:

1. Constructing an adjacency graph: Let's show a graph with G , m nodes. The i -th node corresponds to the a_i data point. There are two ways to create an adjacency graph. [17]:

- K nearest neighbors: Put an edge oriented from node i to j if a_u is among the K nearest neighbors of a_i .
- ϵ neighborhood: u if $\|a_u - a_i\| \leq \epsilon$ and Put an edge between nodes i .

The graphic generated by the first method is a graphic, built in the second method a non-directional graph. It is difficult to choose a good ϵ in many real world applications. That's it, the graph we adopt the knn method to create similarity.

2. Computing the weights: In this step, the weights at the edges have been calculated. Let W denote the weight matrix with W_{iu} having the weight of the edge from node i to node u , and 0 if there is no such edge. The weights on the edges can be calculated by minimizing the following objective function [17],

$$\min \sum_i \|a_i - \sum_u w_{iu} a_u\| \quad (5)$$

with constraints

$$\sum_u w_{iu} = 1, \quad u=1,2,\dots,m \quad (6)$$

3. Computing the Projections: In this step, the linear projections have been calculated. Solve the following generalized eigenvector problem [17]:

$$AKA^T a = \lambda AA^T a \quad (7)$$

where

$$A = (a_1, \dots, a_m)$$

$$K = (Im - W)^T (Im - W)$$

$$Im = \text{diag}(1, \dots, 1) \quad (8)$$

It is easy to check that K is symmetric. Let the column vectors x_0, \dots, x_{d-1} be the solutions of Equation (9), ordered according to their eigenvalues, $\lambda_0 \leq \dots \leq \lambda_{d-1}$. Thus, the embedding is as follows:

$$a_i \rightarrow y_i = A^T a_i$$

$$X = (x_0, \dots, x_{d-1}) \quad (9)$$

where d -dimensional vector is y_i , and X is an $n \times d$ matrix.

E. Random Forest (RF)

RF is a collection of tree type classifiers. The Gini index is used in RF classifier. The division position according to the smallest Gini index is determined by Gini measurements. To generate a tree with the RF classifier, two external parameters must be entered. These parameters are m and N parameters. m is the number of variables used in each node. N is the number of trees that will be developed to determine the best partition [18,19]. The start m value is entered randomly from the outside. Subsequent m 's are reduced or increased relative to the overall error rate. Classification accuracy is understood by generalized error data. P new training data is generated from R training data. Tree-type classifiers in the RF classifier are used in the $\{h(x, \theta_p)P = 1, \dots\}$ type. where, x is the input data; θ_p indicates the random vector. The $h(x, R_p)$ classifier is constructed using the new training data set. The x and y are not found in R_k . When a random pixel is selected for a given training data set R , this pixel belongs to class S_i . Therefore, the Gini index is expressed as below.

$$\sum \sum_{j \neq i} (f(S_i, R)/|R|)(f(S_j, R)/|R|) \quad (10)$$

Here, R is the training data set, S_i is the class to which a randomly selected pixel belongs, $f(S_i, R)/|R|$ indicates the possibility of belonging to the S_i class of the selected example.

In this article has been used the MSER features. The dimensions of these features have been reduced to lower dimensions with help the DFT-LTSA, and DFT-NPE methods by the Linear Discriminant Analysis (LDA) method. These lower size features have been classified by Random Forest (RF) method. The steps of this article are shown in Fig. 1.

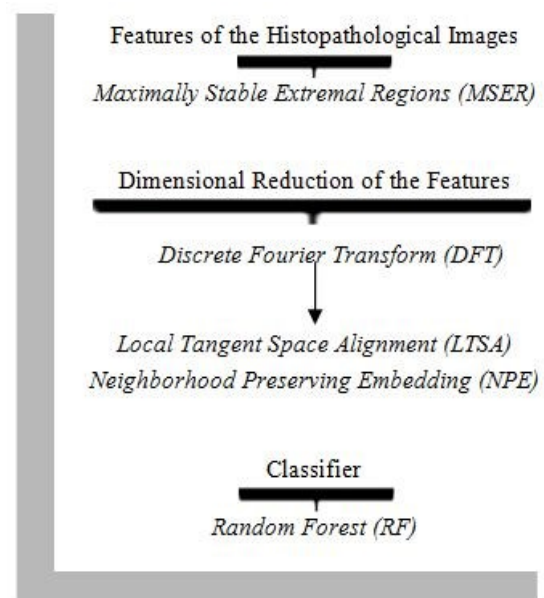


Figure 1. Steps to Apply of the MSER_DFT_LTSA-NPE_RF Algorithms

III. EXPERIMENTAL RESULTS AND DISCUSSION

MSER features of stomach histopathology images were found. The dimensions of the MSER properties for each image are very high and have been obtained differently. It takes a long time to process with MSER characteristics which are different from each other and very high dimensions. By applying the DFT method to the MSER features obtained from this aspect, the cells obtained for the MSER features are converted into vectors, and the same size values are obtained for all the images.

TABLE 1. EXAMPLES OF DIFFERENT DIMENSIONAL CELLS FOUND FOR MSER FEATURES

MSER cells for normal images	MSER cells for benign images	MSER cells for malign images
52800x1	98944x1	66944x1
34368x1	56000x1	70464x1
59776x1	41984x1	74560x1
60544x1	30656x1	81856x1
54016x1	38976x1	73600x1
46016x1	83072x1	65344x1
51712x1	80256x1	60416x1
52480x1	35264x1	54784x1
53184x1	24704x1	51712x1
48192x1	30528x1	55744x1
50432x1	41920x1	81280x1
39488x1	78912x1	124288x1
52544x1	43520x1	80960x1
75904x1	34688x1	112768x1
63872x1	53248x1	121088x1
58304x1	43520x1	94080x1

As shown in Table 1, the number of MSER properties obtained for each image is different from each other. We have applied the DFT method to these obtained property values because the numbers of these properties are different for each image.

When the DFT method is applied, the feature size obtained for normal images is 60x183680. It is 60x197888 for benign images. 60x379008 for malign images. Finally, LTSA and NPE size reduction methods were applied to feature values of these images. Feature values of 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 for LTSA and NPE size reduction methods were discussed respectively. 30 normal + 30 benign + 30 malign images were used for testing purposes while 180 images were used for 30 normal + 30 benign + 30 malign educational purposes. And the test images were classified with the random forest classifier. According to the feature numbers selected in Table 2, the classification performances obtained with the RF classifier are shown. According to Table 2 The highest classification performance obtained with LTSA is 94.44%. This ratio was obtained by selecting 15 properties. However, the highest classification performance obtained with NPE is 86.66%. This ratio was obtained by selecting 10 properties. Based on these results, the LTSA size reduction method over the MSER_DFT property values showed higher classification performance than the NPE size reduction method. As shown in Table 1, both the LTSA and the NPE's classification performance decreases as the number of features increases. This is shown in the graphic in Figure 2.

TABLE 2. CLASSIFICATION PERFORMANCES OF RF CLASSIFIER ACCORDING TO SELECTED FEATURE NUMBER

Classification Performances (%)										
Selected Feature Number	5	10	15	20	25	30	35	40	45	50
MSER_DFT_LTSA_RF	92.22	91.11	94.44	87.77	88.88	82.22	77.77	68.88	54.44	46.66
MSER_DFT-NPE_RF	84.44	86.66	81.11	76.66	66.66	65.55	56.66	50.00	53.33	47.77

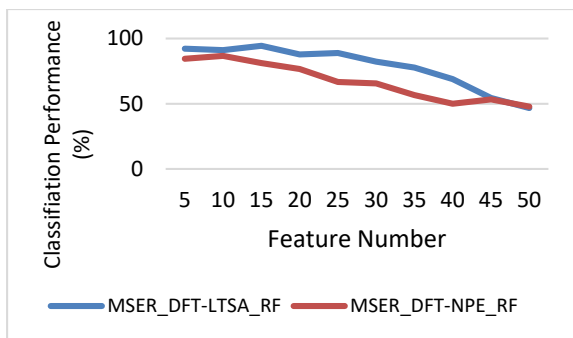


Figure 2. Comparison of the Classification Performances found with MSER_DFT_LTSA_RF and MSER_DFT_NPE_RF Algorithms

In Table 3, GLCM is Gray Level Co-occurrence Matrix Features. LBP is Local Binary Patterns Features. LPP is Locality Preserving Projections for dimensional reduction. HOG is Histograms of Oriented Gradient Feature. LDA is Linear Discriminant Analysis. ANN is Artificial Neural Network.

In Table 3, HOG_LDA_ANN accuracy rate was found as 88,9 %. LBP_LPP_ANN accuracy rate was found as 85.56 %. GLCM_LPP_ANN accuracy rate was found as 80.12%. Our Method MSER_DFT_LTSA_RF accuracy rate has been found as 94.44%.

TABLE 3. COMPARISON OF DIFFERENT RESULTS

Compared	
Method	Accuracy
Our Method MSER_DFT_LTSA_RF	94.44 %
Our Method MSER_DFT_NPE_RF	86.66 %
HOG_LDA_ANN [1]	88,9 %.
GLCM_LPP_ANN [1]	80.12 %
LBP_LPP_ANN [20]	85.56 %

IV. CONCLUSION

Until today, many classification methods have been used in the field of health [21-33]. 180 gastric images taken with Light Microscope help in this article have been classified. Maximally Stable Extremal Regions (MSER) features of the images for classification has been calculated. These MSER features have been applied Discrete Fourier Transform (DFT) method. High-dimensional of these MSER-DFT feature vectors is reduced to lower-dimensional with Local Tangent Space Alignment (LTSA) and Neighborhood Preserving Embedding (NPE). When size reduction process was done, properties in 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 dimensions have been obtained. These low-dimensional data are classified by Random Forest (RF) classification. Thus, MSER_DFT_LTSA-NPE_RF method for gastric histopathological images have been developed. The highest classification performance obtained with LTSA is 94.44%. This ratio was obtained by selecting 15 properties. However, the highest classification performance obtained with NPE is 86.66%. Classification results obtained with these methods have been compared with other classification results in the literature. According to the other methods, our classification results for gastric histopathological images have been seen to be higher. In future studies, an analysis will be performed by applying different feature extraction methods to different cancer images.

Acknowledgements

The authors would like to thank to Prof. Dr. İbrahim Hanifi ÖZERCAN in the pathology department of the Firat University Hospital.

REFERENCES

- [1] KORKMAZ, Sevcin Aytaç, et al. A expert system for stomach cancer images with artificial neural network by using HOG features and linear discriminant analysis: HOG_LDA_ANN. In: Intelligent Systems and Informatics (SISY), 2017 IEEE 15th International Symposium on. IEEE, (2017):000327-000332.
- [2] Lambert, R., Guilloux, A., Oshima, A., Pompe-Kirn, V., Bray, F., Parkin, M., Ajiki, W., Tsukuma, H. (2002). Incidence and mortality from Stomach cancer in Japan, Slovenia and the USA. *International Journal Of Cancer*, 97, 811–818.
- [3] Hirayama, Akiyoshi, et al. "Quantitative metabolome profiling of colon and Stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry." *Cancer research* 69.11 (2009): 4918-4925.
- [4] Brenner, Hermann, Dietrich Rothenbacher, and Volker Arndt. "Epidemiology of Stomach cancer." *Cancer Epidemiology: Modifiable Factors* (2009): 467-477.
- [5] Fujioka, N., Morimoto, Y., Arai, T., Kikuchi, M. (2004). Discrimination between normal and malignant human Stomach tissues by Fourier transform infrared spectroscopy. *Cancer Detection and Prevention*, 28(1), 32–36.
- [6] Yoshihiro, S., Ryukichi, H., Tetsuro Y., Norihiro H., Tatsuya M. et al. (2010). Computer-aided estimation for the risk of development of Stomach cancer by image processing. *Artificial Intelligence in Theory and Practice III*, 331, 197-204.
- [7] Ahmadzadeh, D., Fiuzy, M., Haddadnia, J. (2013). Stomach Cancer Diagnosis by Using a Combination of Image Processing Algorithms, Local Binary Pattern Algorithm and Support Vector Machine. *Journal of Basic and Applied Scientific Research*, 3(2), 243-251.
- [8] Akbari, H., Uto, K., Kosugi, Y., Kojima, K., Tanaka, N. (2011). Cancer detection using infrared hyperspectral imaging. *The Official Journal of The Japanese Cancer Association*, 102(4), 852-857.
- [9] Tannapfel, A., Schmelzer, S., Benicke, M. et al. (2001). Expression of the p53 homologues p63 and p73 in multiple simultaneous Stomach cancer. *The Journal of Pathology*, 195, 163–170.
- [10] Onishi, Katsunori, Tetsuya Takiguchi, and Yasuo Ariki. "3D human posture estimation using the HOG features from monocular image." *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008.
- [11] Matas, J., Chum, O., Urban, M., Pajdla, T., Robust wide-baseline stereo from maximally stable extremal regions, *British Machine Vision Conference*, s.384-393, 2002.
- [12] J. Matas, O. Chum, M. Urban, ve T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, Sayı 22, 761-767, 2004
- [13] Güney M., Arica N.; " Desen Tabanlı İlgi Bölgesi Tespiti" *Journal of Naval Science and Engineering* 2009, Vol. 5 , No.1, pp. 94-106
- [14] Cooley, J.W., Tukey, J.W., An Algorithm for the Machine Computation of Complex Fourier Series, *Mathematics of Computation*, Vol. 19, pp. 297-301, April 1965.
- [15] http://web.itu.edu.tr/~baykut/lab/pdf/Deney_3.pdf
- [16] http://www.math.pku.edu.cn/teachers/yaoy/Spring2011/lecture11_2.pdf
- [17] He, X., Cai, D., Yan, S., & Zhang, H. J. (2005, October). Neighborhood preserving embedding. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on* (Vol. 2, pp. 1208-1213). IEEE.

- [18] Pal, Mahesh. "Random forest classifier for remote sensing classification." *International Journal of Remote Sensing*, vol. 26, no. 1, pp. 217-222, 2005.
- [19] Korkmaz, Sevcen Aytac, and Hamidullah Binol. "Analysis of Molecular Structure Images by using ANN, RF, LBP, HOG, and Size Reduction Methods for early Stomach Cancer Detection." *Journal of Molecular Structure* (2017).
- [20] Korkmaz, S. A. (2018). LBP Özelliklerine Dayanan Lokasyon Koruyan Projeksiyon (LPP) Boyut Azaltma Metodunun Farklı Sınıflandırıcılar Üzerindeki Performanslarının Karşılaştırılması. *Sakarya University Journal of Science*, 22(4), 1-1.
- [21] Korkmaz, S. Aytac, and Mustafa Poyraz. "A New Method Based for Diagnosis of Breast Cancer Cells from Microscopic Images: DWEE--JHT." *Journal of medical systems* 38.9 (2014): 1.
- [22] Korkmaz, Sevcen Aytac, and Mustafa Poyraz. "Least square support vector machine and minimum redundancy maximum relevance for diagnosis of breast cancer from breast microscopic images." *Procedia-Social and Behavioral Sciences* 174 (2015): 4026-4031.
- [23] Korkmaz, Sevcen Aytac, and Mehmet Fatih Korkmaz. "A new method based cancer detection in mammogram textures by finding feature weights and using Kullback-Leibler measure with kernel estimation." *Optik-International Journal for Light and Electron Optics* 126.20 (2015): 2576-2583.
- [24] Korkmaz, Sevcen Aytac, Mehmet Fatih Korkmaz, and Mustafa Poyraz. "Diagnosis of breast cancer in light microscopic and mammographic images textures using relative entropy via kernel estimation." *Medical & biological engineering & computing* 54.4 (2016): 561-573.
- [25] KORKMAZ, Sevcen Aytac; EREN, Haluk. Cancer detection in mammograms estimating feature weights via Kullback-Leibler measure. In: *Image and Signal Processing (CISP), 2013 6th International Congress on. IEEE*, 2 (2013):1035-1040.
- [26] KORKMAZ, Sevcen AYTAÇ. "DETECTING CELLS USING IMAGE SEGMENTATION OF THE CERVICAL CANCER IMAGES TAKEN FROM SCANNING ELECTRON MICROSCOPE." *The Online Journal of Science and Technology-October* 7.4 (2017).
- [27] KORKMAZ, Sevcen Aytac, et al. Recognition of the stomach cancer images with probabilistic HOG feature vector histograms by using HOG features. In: *Intelligent Systems and Informatics (SISY), 2017 IEEE 15th International Symposium on. IEEE*, (2017). p. 000339-000342.
- [28] Korkmaz, Sevcen Aytac, et al. "Diagnosis of breast cancer nano-biomechanics images taken from atomic force microscope." *Journal of Nanoelectronics and Optoelectronics* 11.4 (2016): 551-559.
- [29] Korkmaz, S. A., Poyraz, M., Bal, A., Binol, H., Özercan, I. H., Korkmaz, M. F., & Aydin, A. M. (2015). New methods based on mRMR_LSSVM and mRMR_KNN for diagnosis of breast cancer from microscopic and mammography images of some patients. *International Journal of Biomedical Engineering and Technology*, 19(2), 105-117.
- [30] Sengur, Abdulkadir. "Support vector machine ensembles for intelligent diagnosis of valvular heart disease." *Journal of medical systems* 36.4 (2012): 2649-2655.
- [31] A. Şengür, "An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnostic of valvular heart diseases" *Computers in Biology and Medicine* 38 (2008) 329 – 338.
- [32] Sengur, Abdulkadir, Ibrahim Turkoglu, and M. Cevdet Ince. "Wavelet packet neural networks for texture classification." *Expert systems with applications* 32.2 (2007): 527-533.
- [33] Sengur, Abdulkadir. "Multiclass least-squares support vector machines for analog modulation classification." *Expert Systems with Applications* 36.3 (2009): 6681-6685.